

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : <b>G01N 27/00, 27/26, 33/543, 33/551, C12Q 1/68, A01N 1/02, C12M 1/00</b>		<b>A1</b>	(11) International Publication Number: <b>WO 98/08083</b> (43) International Publication Date: 26 February 1998 (26.02.98)
(21) International Application Number: <b>PCT/US97/14372</b> (22) International Filing Date: 13 August 1997 (13.08.97) (30) Priority Data: 08/700,182 20 August 1996 (20.08.96) <b>US</b> (71) Applicant (for all designated States except US): <b>MOTOROLA INC. [US/US]; 1303 East Algonquin Road, Schaumburg, IL 60196 (US).</b> (72) Inventors; and (75) Inventors/Applicants (for US only): <b>REBER, William, L. [US/US]; 1029 Buccaneer Road #6, Schaumburg, IL 60916 (US). ACKLEY, Donald, E. [US/US]; 2033 Cambridge Avenue, Cardiff, CA 92007 (US). PERTTUNEN, Cary, D. [US/US]; 11764 Raintree Court, Shelby Township, MI 48315 (US).</b> (74) Agents: <b>SARLI, Anthony, J., Jr. et al.; Motorola Inc., Intellectual Property Dept., 1303 East Algonquin Road, Schaumburg, IL 60196 (US).</b>			(81) Designated States: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</b>  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: <b>METHOD AND APPARATUS FOR DETECTING PREDETERMINED MOLECULAR STRUCTURES IN A SAMPLE</b>			
(57) Abstract <p>A predetermined molecular structure in a sample is detected by sensing a pattern in which the sample binds to an array of binding sites in a molecular detection device (20), and comparing the pattern to a reference pattern to detect the predetermined molecular structure in the sample (22). In one embodiment, the reference pattern is generated by sensing a pattern in which a reference sample containing the predetermined molecular structure binds to a like array of binding sites. In another embodiment, the reference pattern is generated by predicting a pattern in which the predetermined molecular structure binds to the array of binding sites.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

5 METHOD AND APPARATUS FOR DETECTING PREDETERMINED  
MOLECULAR STRUCTURES IN A SAMPLE

Field of the Invention

10 The present invention relates to methods and  
system for molecular detection.

Background of the Invention

15 An increased effort has been directed toward  
the development of chips for molecular detection.  
Typically, a molecular detection chip includes a  
substrate on which an array of binding sites is  
arranged. Each binding site, or hybridization  
20 site, has a respective molecular receptor which  
binds or hybridizes with a molecule having a  
predetermined structure.

A sample solution is applied to the molecular  
detection chip, and molecules in the sample bind or  
hybridize at one or more of the binding sites. The  
25 particular binding sites at which hybridization  
occurs are detected, and one or more molecular  
structures within the sample are subsequently  
deduced.

30 Of great interest are molecular detection  
chips for gene sequencing. These chips, often  
referred to as DNA chips, utilize an array of  
selective binding sites each having respective  
single-stranded DNA probes. A sample of single-  
35 stranded DNA fragments, referred to as target DNA,  
is applied to the DNA chip. The DNA fragments  
attach to one or more of the DNA probes by a  
hybridization process. By detecting which DNA  
probes have a DNA fragment hybridized thereto, a

5 sequence of nucleotide bases within the DNA fragment can be determined.

One use of molecular detection chips is to perform medical diagnostics. Here, a genomic sample from an individual is screened to determine  
10 if the individual has a genetically-inherited disease or a genetic predisposition to disease.

As medical diagnostics migrate to miniature devices capable of performing higher numbers of tests with greater sensitivity, there is a  
15 corresponding increase in the amount of information being generated, and in the demand for greater processing power to obtain test results.

#### Brief Description of the Drawings

20 The invention is pointed out with particularity in the appended claims. However, other features of the invention will become more apparent and the invention will be best understood  
25 by referring to the following detailed description in conjunction with the accompanying drawings in which:

FIG. 1 is a flow chart of an embodiment of a method of detecting a predetermined molecular structure in a sample;  
30

FIG. 2 is a flow chart of an embodiment of a method of generating the reference pattern;

FIG. 3 is a flow chart of another embodiment of a method of generating the reference pattern;

35 FIG. 4 illustrates an array of binding sites on a molecular detection device used in the specific embodiment;

FIG. 5 illustrates a reference pattern for detecting an a-c-g nucleotide sequence in a sample;

5           FIG. 6 illustrates a reference pattern for detecting an a-c-t nucleotide sequence in a sample;

          FIG. 7 is an example of a pattern generated by a sample having an unknown molecular structure;

10           FIG. 8 is a flow chart of additional steps which can be utilized to detect the predetermined molecular structure;

          FIG. 9 is a block diagram of an apparatus for detecting a predetermined molecular structure in a sample; and

15           FIG. 10 is a flow chart of an embodiment of a method of gene discovery in accordance with the present invention.

#### Detailed Description of a Preferred Embodiment

20

          Embodiments of the present invention advantageously provide improved information processing approaches to detecting predetermined molecular structures using a miniaturized device having an array of biological sensors. Just as semiconductor devices are designed to perform specific functions, a diagnostic device in accordance with the present invention is designed to perform one or more specific diagnostic tests.

25

30           FIG. 1 is a flow chart of an embodiment of a method of detecting a predetermined molecular structure in a sample. In general, the method can be utilized for detection of a variety of molecular structures in a variety of different types of samples. Examples of the different types of samples include, but are not limited to, medical samples, environmental samples, agricultural samples, and other samples applicable to

35

diagnostics.

5           The predetermined molecular structure can be  
any indication of a structure of molecules  
contained in the sample. For example, the  
predetermined molecular structure can be indicative  
10 of a presence of a pathogen in an environmental  
sample such as water. In an agricultural  
application, the predetermined molecular structure  
can provide an indication of crop resistance, for  
example. In a medical application, a predetermined  
15 molecular structure can be indicative of a disease  
gene.

Of particular interest is the detection of a  
predetermined nucleotide base sequence in a sample  
from a living organism or a plant. For a sample  
obtained from an individual, the sample can include  
20 a DNA sample or an RNA sample, for example. The  
predetermined nucleotide base sequence can be  
associated with a genetically-inherited disease  
and/or a genetic predisposition to disease of the  
individual, for example. Detection of these types  
25 of molecular structures allow medical personnel to  
formulate an appropriate treatment strategy for the  
individual.

As indicated by block 10, a step of providing  
a molecular detection device is performed. Various  
30 types of molecular detection devices can be  
utilized. Such molecular detection devices  
include, but are not limited to, molecular  
detection chips, DNA chips, biosensor arrays,  
genosensor arrays, and the like.

35           The molecular detection device typically  
includes a plurality binding sites or hybridization  
sites which are arranged as an array on a  
substrate. Each binding site has a respective  
molecular receptor which specifically binds or

5 hybridizes with a molecule having a predetermined structure.

Each molecular receptor typically includes a biological or synthetic molecule having a specific affinity to the molecule to be detected. Of particular interest is a molecular receptor having a chain of at least one nucleotide to hybridize with a molecule having a complementary chain of at least one nucleotide. Here, for example, the molecular receptor can include a DNA probe for detecting a corresponding, complementary DNA sequence in the sample.

It is noted, however, that the scope of the invention is not limited to sensing the hybridization of DNA molecules. For example, embodiments of the present invention can be utilized to detect RNA hybridization and antibody-antigen binding events. As another alternative, the molecular detection device can include an array of detection sites, such as in the context of an oligonucleotide ligation assay (OLA). Using a ligase chain reaction, pairs of oligonucleotides are utilized to amplify a selected oligonucleotide sequence. To detect the selected oligonucleotide sequence, a corresponding detection site is screened for full-length ligated oligonucleotides using any of the sensing approaches described herein.

As indicated by block 12, an optional step of tagging molecules within the sample is performed. Each molecule is tagged with a member which can be sensed by the molecular detection device. Such members are commonly referred to in the art as tags, markers, and labels. Examples of such members include, but are not limited to,

5 radioactive members, optical members (such as  
fluorescent members, luminescent members, and  
light-scattering members), charged members, and  
magnetic members.

10 As indicated by block 14, a step of applying  
the sample to the molecular detection device is  
performed. Thereafter, the sample is allowed to  
hybridize at one or more of the binding sites, as  
indicated by block 16. Typically, the sample  
specifically binds to at least one of the binding  
15 sites, and non-specifically binds to at least  
another one of the binding sites. By specific  
binding, it is meant that an intended target  
molecule is bound to a molecular receptor. By non-  
specific binding, it is meant that an unintended  
20 target molecule is bound to a molecular receptor.  
In the case where the molecular receptor includes a  
chain of at least one nucleotide, specific binding  
occurs with a molecule having a complementary chain  
of the at least one nucleotide. Non-specific  
25 binding with such a molecular receptor occurs with  
a molecule having at least one mismatching base.

To hasten the hybridization process, a local  
concentration of molecules in the sample can be  
increased at predetermined sites using electric  
30 field enhancements. Here, each site has an  
electrode associated therewith for selectively  
generating an electric field thereby. The electric  
field is generated by applying an electric  
potential between an electrode at the site and a  
35 counter electrode. To attract molecules to the  
site, the polarity of the electric potential is  
selected to generate an electric field having a  
polarity opposite to the charge of the molecules.



5           After hybridization, an optional step of  
removing unwanted molecules from the binding sites  
can be performed, as indicated by block 18. The  
step of removing unwanted molecules can be  
performed by generating an electric field having  
10           the same polarity as the charge of the unwanted  
molecules. The electric field acts to repel  
unwanted molecules from the binding sites. As an  
alternative to, or in conjunction with, the field-  
based approach, a thermally-assisted approach can  
15           be utilized to remove unwanted molecules. Here,  
the temperature at the binding sites is raised, in  
dependence upon a melting temperature, to  
dissociate partially-bound molecules from the  
molecular receptors. Regardless of the approach  
20           utilized, the unwanted molecules to be dehybridized  
can include unbound molecules and partially-bound  
(i.e. non-specifically bound) molecules.

          Typically, the step of removing unwanted  
molecules does not remove all unwanted molecules  
25           from the binding sites. This step is beneficial,  
however, in improving the accuracy of detection as  
outlined in subsequent steps.

          As indicated by block 20, the method includes  
a step of sensing a pattern in which the sample  
30           binds to an array of binding sites in a molecular  
detection device. The pattern can be sensed using  
a variety of approaches, including but not limited  
to, optical approaches, radioactive-sensing  
approaches, electronic approaches, and magnetic  
35           approaches. The specific approach utilized depends  
upon the type of tagging member attached to the  
molecules in the sample.

          Preferably, the step of sensing the pattern  
includes sensing an intensity or a magnitude of

5 binding at each of a plurality of the array of  
binding sites. Each intensity can be indicative of  
a number of molecules bound at a respective binding  
site, a binding strength at a respective binding  
10 site, and/or a melting temperature for the binding  
at a respective binding site.

For the purpose of illustration, the method  
will be described using an optical sensing  
approach. Here, the step of sensing the pattern  
includes capturing at least one image of the  
15 pattern. Typically, the image is formed by  
luminescent light, fluorescent light or a  
scattering of light associated with a hybridization  
event. The image can be captured using a CCD  
camera, a confocal microscope, or other like  
20 imaging device.

Preferably, the image indicates an intensity  
or magnitude of binding at each of the binding  
sites by a measure of illumination. The measure of  
illumination at a binding site can be based on a  
25 gray scale or the like at the location of the  
binding site on the image. In one embodiment, the  
intensity of binding at a binding site is based on  
an average illumination (e.g. an average gray  
scale) at the binding site.

30 As indicated by block 22, the method includes  
a step of comparing the pattern to a reference  
pattern to detect the predetermined molecular  
structure in the sample. In one embodiment, the  
step of comparing includes determining a  
35 correlation between the pattern and the reference  
pattern. Here, the predetermined molecular  
structure can be detected when a measure of the  
correlation is within a predetermined range. In  
another embodiment, the step of comparing includes

5 determining a difference between the pattern and the reference pattern. Here, the predetermined molecular structure can be detected when a measure of the difference is within a predetermined range.

10 For optical sensing embodiments, the step of comparing includes a step of comparing at least one image of the pattern to an image of the reference pattern.

15 As indicated by block 24, the method optionally comprises a step of determining a confidence level of detecting the predetermined molecular structure in the sample. The confidence level indicates a degree of significance of the result obtained in the step of comparing in block 22.

20 To screen the sample for a plurality of different molecular structures, the steps indicated by blocks 22 and 24 can be repeated for a plurality of different reference patterns. Here, for example, a genomic sample can be screened to  
25 determine if it contains any of a plurality of predetermined base sequences.

FIG. 2 is a flow chart of an embodiment of a method of generating the reference pattern. Typically, the reference pattern is generated prior  
30 to performing the steps indicated in FIG. 1.

As indicated by block 30, the method includes a step of providing a reference device having a like array of binding sites as the molecular  
35 detection device used for detection. If desired, the same molecular detection device can be utilized for generating the reference pattern and for subsequent detection of an unknown molecular structure in a sample. Typically, however, another like device is utilized.

5           As indicated by block 32, an optional step of tagging molecules in a reference sample is performed. The reference sample is selected to contain the predetermined molecular structure which is to be detected.

10           As indicated by block 34, a step of applying the reference sample to the reference device is performed. As indicated by block 36, the reference sample is allowed to hybridize at one or more of the binding sites. Optionally, a step of removing  
15           unwanted molecules from the binding sites is performed, as indicated by block 38.

          As indicated by block 40, a step of sensing a pattern in which the reference sample binds to the like array of binding sites is performed. This  
20           pattern is utilized as the reference pattern for subsequent comparison. Preferably, the reference pattern is sensed using the same approach as utilized for subsequent detection steps. It is  
25           further preferred that an intensity of binding at each of the binding sites in the like array be included in the reference pattern.

          As indicated by block 42, an optional step of modifying a temperature at the binding sites is performed. Preferably, the temperature is raised  
30           to dissociate weakly bound molecules from the binding sites. Thereafter, the step of sensing the pattern, indicated by block 40, is repeated. The temperature can be repeatedly raised to generate a plurality of temperature-dependent reference  
35           patterns.

          Thereafter, the reference device can be washed, and the steps indicated by blocks 32, 34, 36, 38, 40, and 42 can be repeated for another reference sample containing the same predetermined

5 molecular structure. Alternatively, the steps indicated by blocks 30, 32, 34, 36, 38, 40, and 42 can be repeated to apply the same predetermined molecular structure to a number of like reference devices. Either approach may be utilized to  
10 provide a plurality of reference patterns for the same predetermined molecular structure.

Sensed patterns formed by a sample having an unknown molecular structure can be compared to each of the above-described plurality of reference patterns, or a statistical model thereof, to detect  
15 the predetermined molecular structure in the sample.

FIG. 3 is a flow chart of another embodiment of a method of generating the reference pattern. As indicated by block 50, the method includes a  
20 step of determining an architecture of the array of binding sites of the molecular detection device. This step can include determining a layout of the binding sites, and determining the type of  
25 molecular receptor at each of the binding sites.

As indicated by block 52, the method includes a step of predicting a reference pattern in which the predetermined molecular structure binds to the array of binding sites. The reference pattern is  
30 predicted based upon the predetermined molecular structure and the architecture of the molecular detection device. Preferably, the reference pattern includes a predicted intensity of binding at each of the binding sites.

35 Regardless of the approach taken, the reference pattern acts as a novelty filter which is predictive of a successful or a desirable test result.

5           FIGS. 4 to 7 illustrate the detection of a  
predetermined molecular structure in accordance  
with a specific embodiment of the present  
invention. In this simplified example, a  
predetermined sequence of nucleotide bases is  
10 detected in a genomic sample, such as a DNA sample.

FIG. 4 illustrates an array of binding sites  
60 on a molecular detection device used in the  
specific embodiment. The array of binding sites 60  
contains all possible single-stranded DNA probes  
15 having a length of three bases. It is noted that  
molecular detection devices used in practice  
typically have more binding sites. For example,  
256x256 arrays containing all possible 8-mer DNA  
probes are commonly utilized.

20           The respective base sequence at each binding  
site is indicated using standard nucleotide  
abbreviations ("a" representing adenine, "c"  
representing cytosine, "g" representing guanine,  
and "t" representing thymine). As is known in the  
25 art, each DNA probe is utilized to detect molecules  
having a complementary sequence of nucleotide  
bases.

FIG. 5 illustrates a reference pattern for  
detecting an a-c-g nucleotide sequence in a sample.  
30 The reference pattern indicates an intensity of  
binding at each of the binding sites by an  
intensity of illumination. Here, darker sites are  
indicative of lower binding intensities, while  
brighter sites are indicative of higher binding  
intensities.

35           A binding site 62 having a t-g-c sequence,  
which is complementary to the a-c-g sequence, has  
the greatest intensity of binding. Binding sites  
having single-mismatching complementary bases, such

5 as binding sites indicated by reference numeral 64,  
have a lesser intensity of binding. Those binding  
sites having two mismatching complementary bases,  
such as those indicated by reference numeral 66,  
have an even lesser intensity of binding. Binding  
10 sites with no matching complementary bases, such as  
those indicated by reference numeral 68, have a low  
intensity of binding.

FIG. 6 illustrates a reference pattern for  
detecting an a-c-t nucleotide sequence in a sample.  
15 The reference patterns in FIGS. 5 and 6 can be  
sensed using a reference sample, or can be  
predicted based on the number of mismatching bases  
at each binding site.

For purposes of illustration, the sequences in  
20 this example are assumed to have a specific  
orientation. As a result, an a-c-t sequence and a  
t-c-a sequence do not specifically hybridize at the  
same binding site. It is noted, however, that this  
should not be construed as a limitation in the  
25 scope of the present invention.

FIG. 7 is an example of a pattern generated by  
a sample having an unknown molecular structure.  
The pattern is generated by applying a sample of  
tagged single-stranded DNA molecules to the  
30 molecular detection device, allowing the molecules  
to hybridize to the binding sites, and optionally  
removing unwanted molecules.

The resulting pattern shows a high intensity  
of binding at a t-g-a site 70. If standard  
35 detection techniques were utilized, one would  
conclude that the sample includes an a-c-t  
nucleotide sequence (i.e. the complement of t-g-a).  
However, the overall pattern is better correlated  
to the reference pattern for the a-c-g nucleotide

5 sequence (in FIG. 5) than to the reference pattern  
for a-c-t (in FIG. 6). Moreover, the pattern is  
significantly correlated to the reference pattern  
in FIG. 5. Hence, the method of the present  
invention would conclude that the sample includes  
10 the a-c-g nucleotide sequence.

The pattern in FIG. 7 can be compared to the  
reference patterns in FIG. 5 and FIG. 6 using any  
of a variety of correlation measures and/or  
difference measures. For example, a difference  
15 measure can be formulated using intensity  
differences (between the pattern and the reference  
pattern) for each of the binding sites. The  
difference measure can be a sum of absolute  
differences, a maximum difference, a root-mean  
20 square difference, or any other suitable function  
of the intensity differences. A correlation  
measure can be formulated using a coefficient of  
correlation between the pattern intensities and the  
reference intensities. In general, any linear or  
25 nonlinear function of the pattern intensities and  
the reference intensities can be utilized to  
compare the pattern to the reference pattern.

FIG. 8 is a flow chart of additional steps  
which can be utilized to detect the predetermined  
30 molecular structure. The steps illustrated here  
can be performed after performing the steps  
indicated by blocks 10, 12, 14, 16 and 18 in FIG.  
1.

As indicated by block 80, a step of modifying  
35 a temperature at the binding sites is performed.  
Preferably, the temperature is raised to dissociate  
weakly bound molecules from the binding sites. The  
temperature can be independently modified at



5 selected binding sites, or can be modified for all of the binding sites.

10 Thereafter, a step of sensing a pattern, indicated by block 82, is performed. By repeatedly raising the temperature and sensing a resulting pattern, a plurality of temperature-dependent patterns is generated.

15 As indicated by block 84, a step of comparing at least one of the plurality of temperature-dependent patterns to a corresponding at least one of a plurality of temperature-dependent reference patterns is performed. Here, for example, each of the temperature-dependent patterns can be compared to a corresponding one of the temperature-dependent reference patterns. Alternatively, only selected  
20 ones of the temperature-dependent patterns can be compared to corresponding reference patterns. A correlation measure and/or a difference measure is computed based on this comparison. A predetermined molecular structure is detected when the measure is within a predetermined range.

25 In one embodiment, the temperature-dependent pattern having a greatest variability of intensity is selected for comparison. The variability of intensity is greatest at a temperature which  
30 dissociates many non-specifically-bound molecules, but does not significantly dissociate specifically-bound molecules. This pattern can be compared with a corresponding reference pattern to detect a predetermined molecular structure. It is noted  
35 that a variety of different measures of variability can be utilized, including but not limited to, sample variance and sample standard deviation.

FIG. 9 is a block diagram of an apparatus for detecting a predetermined molecular structure in a

5 sample. The apparatus includes an array of binding sites 90 each having a respective molecular receptor. A sensor 92 provides means for sensing the pattern in which the sample binds to the array of binding sites 90. Various types of sensors can  
10 be utilized, as described earlier.

A memory 94 contains a library of at least one reference pattern. Preferably, the memory 94 contains a library of a plurality of reference patterns so that the apparatus can be utilized for  
15 detecting any of a plurality of predetermined molecular structures in the sample.

A processor 96 is operatively associated with the sensor 92 and the memory 94. The processor 96 provides means for comparing the pattern sensed by  
20 the sensor 92 to the reference patterns stored in the memory. Based on the comparison(s), the processor 96 outputs an indication of which predetermined molecular structure is detected, and a confidence level therefor.

25 In one embodiment, the array of binding sites 90, the sensor 92, the memory 94, and the processor 96 are integrated to form a single device. Alternatively, selected ones of the above-described components can be external to the device.

30 Regardless of its form, the above-described apparatus can be utilized to perform the steps described for FIGS. 1 and 8. To modify the temperature at the binding sites, one or more heating elements 98 can be incorporated in the  
35 device. The heating elements 98 are preferably controlled by the processor 96 to form a plurality of temperature-dependent patterns.

FIG. 10 is a flow chart of an embodiment of a method of gene discovery in accordance with the

5 present invention. As indicated by block 100, the  
method includes a step of sensing a pattern of  
detection for a sample applied to a molecular  
detection device having a plurality of detection  
10 sites. The sample is taken from a first species  
having unknown genes to be discovered. Any of the  
various approaches described herein can be utilized  
for sensing the pattern.

As indicated by block 12, the method includes  
a step of determining an architecture of the  
15 plurality of detection sites of the molecular  
detection device.

A step of reading a nucleotide sequence from a  
database is performed, as indicated by block 104.  
In general, any nucleotide sequence can be read.  
20 Of particular interest, however, is a nucleotide  
sequence from a second species other than the first  
species. For example, the nucleotide sequence can  
include a gene from a fruit fly, while the sample  
in which gene discovery is to be performed is from  
25 a human.

As indicated by block 106, the method includes  
a step of predicting a reference pattern which  
would be detected if the nucleotide sequence were  
applied to the molecular detection device. As  
30 indicated by block 108, a step of comparing the  
pattern to the reference pattern is performed to  
determine whether the nucleotide sequence is within  
the sample.

The steps of reading a nucleotide sequence  
35 from the database, predicting a reference pattern  
for the nucleotide sequence, and comparing the  
pattern to the reference pattern are repeated to  
discover the presence of genes across different  
species.

5           Thus, there has been described herein a concept, as well as several embodiments including preferred embodiments of a method and apparatus for detecting predetermined molecular structures in a sample.

10           Because the various embodiments of the present invention compare actual test patterns to predicted or sensed reference patterns, they provide a significant improvement in that the results of a diagnostic test can be rapidly determined.

15           Additionally, the various embodiments of the present invention as herein-described use a plurality of temperature-dependent patterns to improve the sensitivity and accuracy of detection.

20           It will be apparent to those skilled in the art that the disclosed invention may be modified in numerous ways and may assume many embodiments other than the preferred form specifically set out and described above.

25           Accordingly, it is intended by the appended claims to cover all modifications of the invention which fall within the true spirit and scope of the invention.

## Claims

5

1. A method of detecting a predetermined molecular structure in a sample, the method comprising the steps of:

10

sensing a pattern in which the sample binds to an array of binding sites in a molecular detection device; and

15

comparing the pattern to a reference pattern to detect the predetermined molecular structure in the sample.

20

2. The method of claim 1 further comprising the step of determining a confidence level of detecting the predetermined molecular structure in the sample.

25

3. The method of claim 1 wherein the step of comparing includes at least one of determining a difference between the pattern and the reference pattern, determining a correlation between the pattern and the reference pattern, comparing at least one image of the pattern to an image of the reference pattern, comparing a plurality of temperature-dependent patterns to a plurality of temperature-dependent reference patterns.

30

35

4. The method of claim 1 wherein an intensity level of at least one of the binding sites is indicative of at least one of molecules bound at a respective binding site, a binding strength at a respective binding site, and a melting temperature at a respective binding site.

5           5. The method of claim 1 wherein the sample includes at least one of an environmental sample and an agriculture sample.

10           6. An apparatus for detecting a predetermined molecular structure in a sample, the apparatus comprising:

          a sensor for sensing a pattern in which the sample binds to an array of binding sites;

          a memory which contains a reference pattern for the predetermined molecular structure; and

15           a processor operatively associated with the sensor and the memory, the processor for comparing the pattern to the reference pattern to detect the predetermined molecular structure in the sample.

20           7. The apparatus of claim 6 wherein the sensor senses the pattern by capturing at least one image of the pattern.

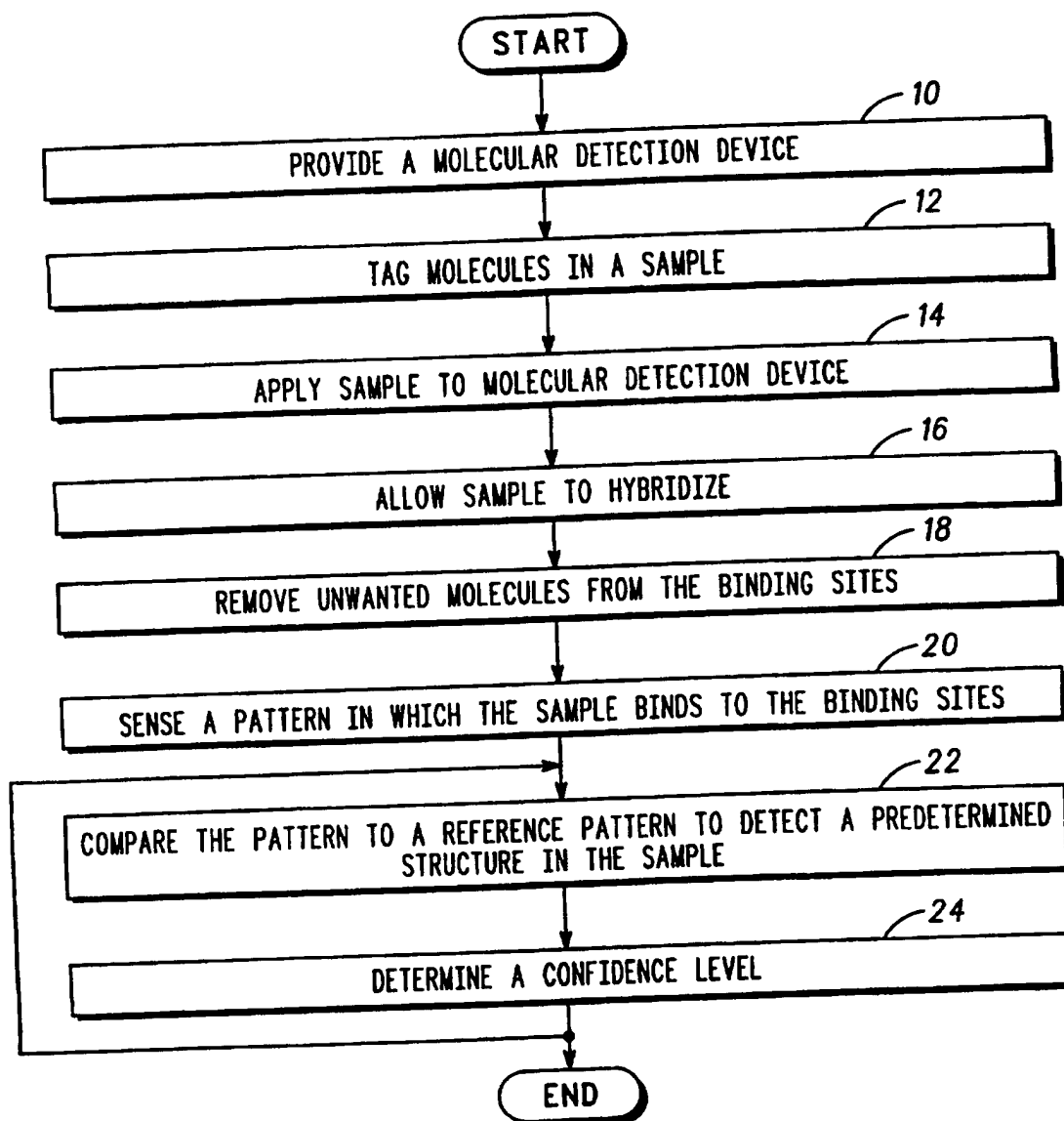
25           8. The apparatus of claim 6 wherein the sensor senses a plurality of temperature-dependent patterns.

30           9. The apparatus of claim 8 wherein the processor selects a temperature-dependent pattern having a greatest variability of intensity, and compares the temperature-dependent pattern having the greatest variability of intensity to the reference pattern.

35           10. The apparatus of claim 6 wherein the reference pattern is generated by sensing a pattern in which a reference sample containing the predetermined molecular structure binds to a like

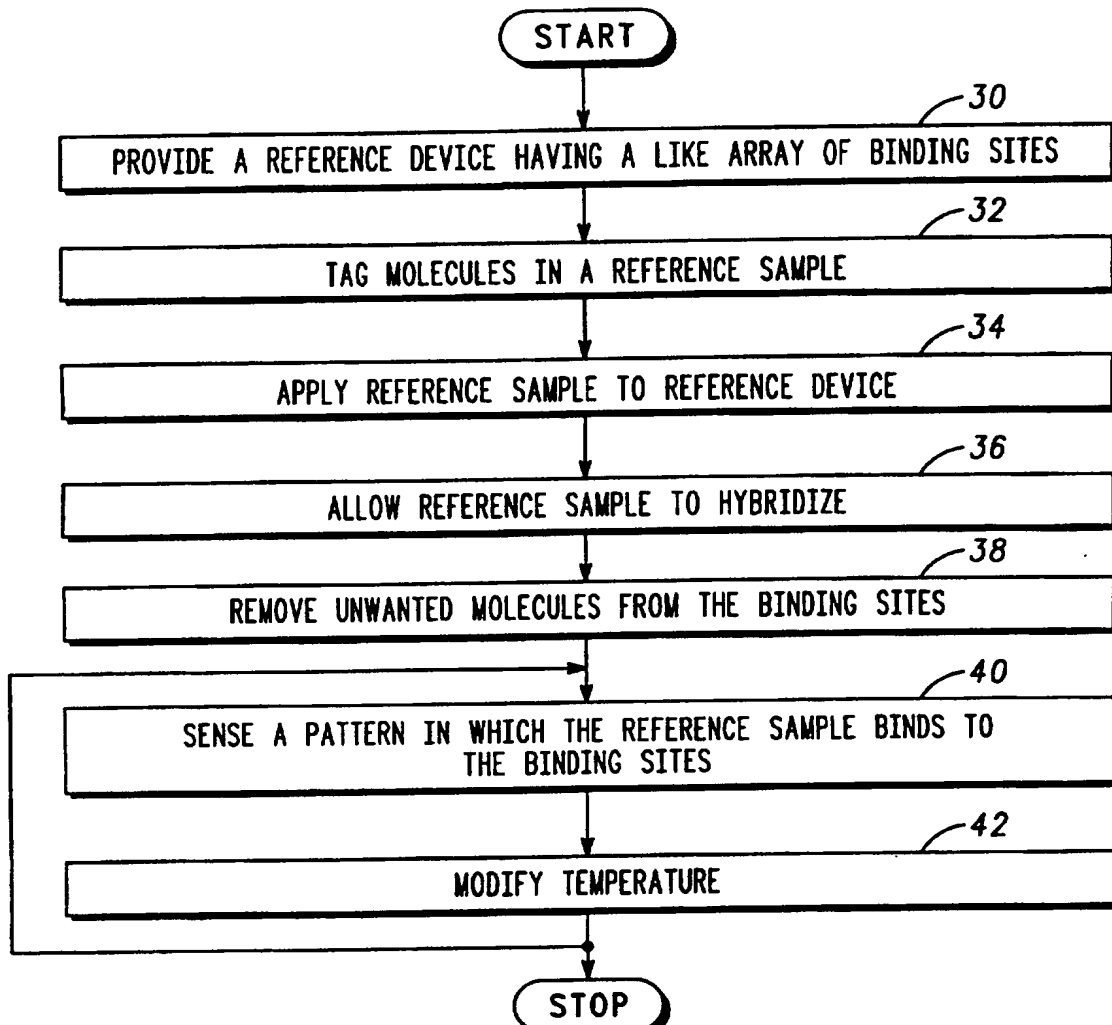
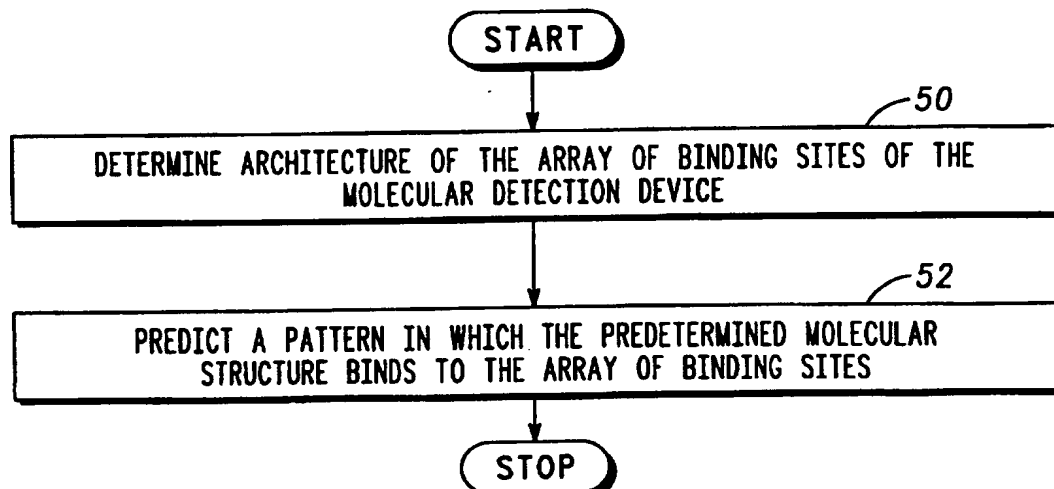
- 5 array of binding sites or by predicting a pattern in which the predetermined molecular structure binds to the array of binding sites.

1/5

**FIG. 1**



2/5

**FIG. 2****FIG. 3**

3/5

AAA	AAC	ACA	ACA	CAA	CAC	CCA	CCC	60
AAG	AAT	ACG	ACT	CAG	CAT	CCG	CCT	
AGA	AGC	ATA	ATC	CGA	CGC	CTA	CTC	
AGG	AGT	ATG	ATT	CGG	CGT	CTG	CTT	
GAA	GAC	GCA	GCC	TAA	TAC	TCA	TCC	
GAG	GAT	GCG	GCT	TAG	TAT	TCG	TCT	
GGA	GGC	GTA	GTC	TGA	TGC	TTA	TTC	
GGG	GGT	GTG	GTT	TGG	TGT	TTG	TTT	

FIG. 4

AAA	AAC	ACA	ACA	CAA	CAC	CCA	CCC	64
AAG	AAT	ACG	ACT	CAG	CAT	CCG	CCT	
AGA	AGC	ATA	ATC	CGA	CGC	CTA	CTC	
AGG	AGT	ATG	ATT	CGG	CGT	CTG	CTT	
GAA	GAC	GCA	GCC	TAA	TAC	TCA	TCC	
GAG	GAT	GCG	GCT	TAG	TAT	TCG	TCT	
GGA	GGC	GTA	GTC	TGA	TGC	TTA	TTC	
GGG	GGT	GTG	GTT	TGG	TGT	TTG	TTT	
68	68	68	64	62	66			

FIG. 5

4/5

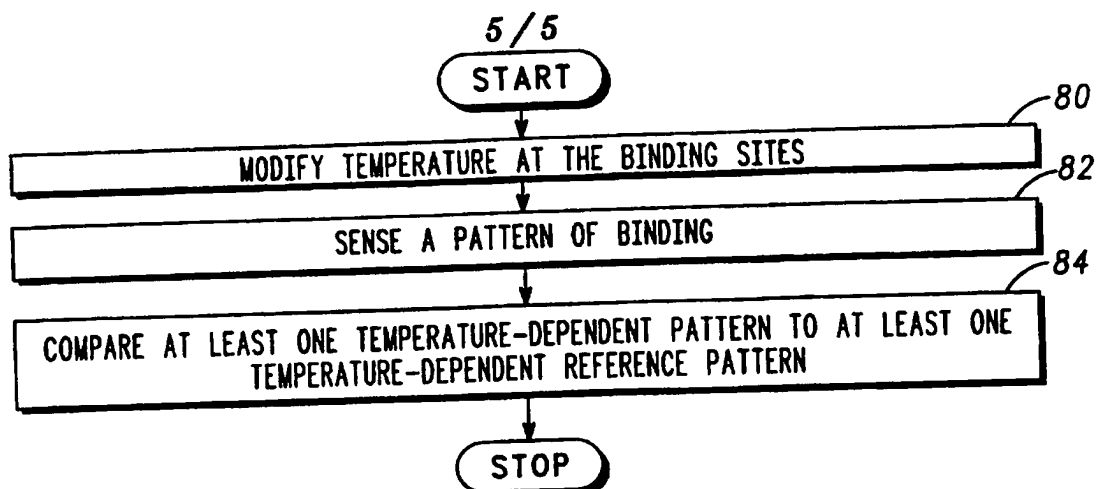
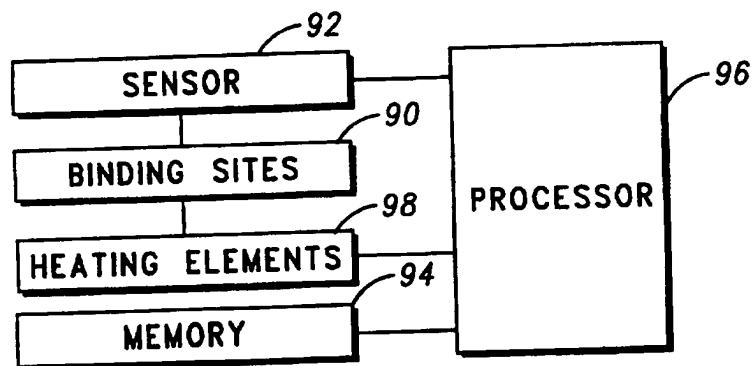
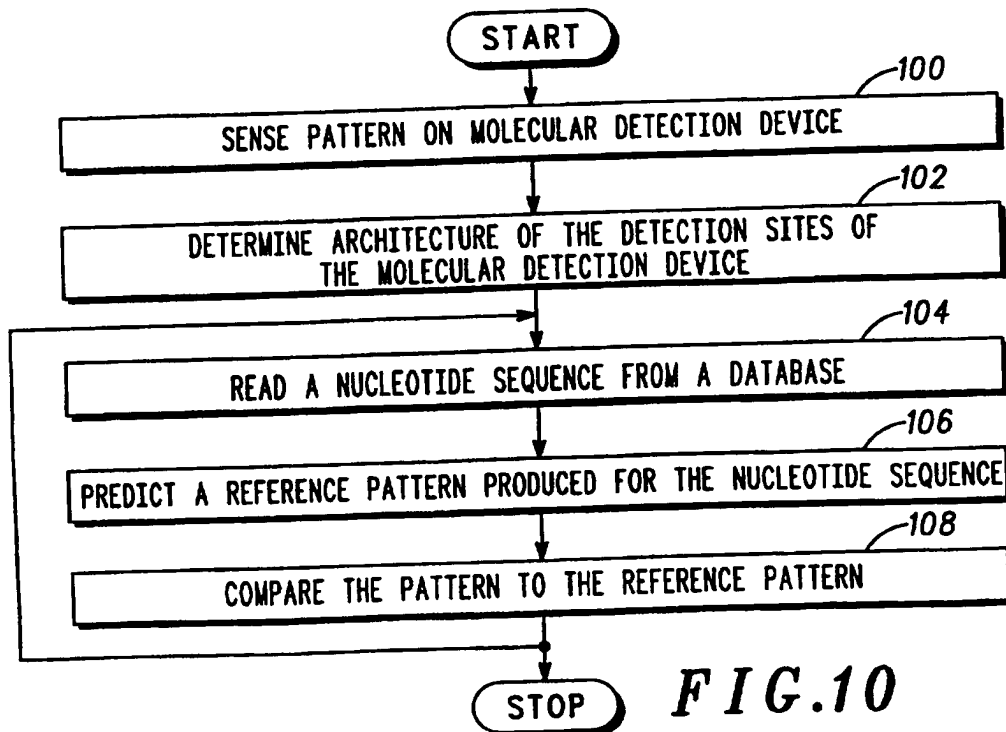
AAA	AAC	ACA	ACA	CAA	CAC	CCA	CCC
AAG	AAT	ACG	ACT	CAG	CAT	CCG	CCT
AGA	AGC	ATA	ATC	CGA	CGC	CTA	CTC
AGG	AGT	ATG	ATT	CGG	CGT	CTG	CTT
GAA	GAC	GCA	GCC	TAA	TAC	TCA	TCC
GAG	GAT	GCG	GCT	TAG	TAT	TCG	TCT
GGA	GGC	GTA	GTC	TGA	TGC	TTA	TTC
GGG	GGT	GTC	GTT	TGG	TGT	TTG	TTT

**FIG. 6**

AAA	AAC	ACA	ACA	CAA	CAC	CCA	CCC
AAG	AAT	ACG	ACT	CAG	CAT	CCG	CCT
AGA	AGC	ATA	ATC	CGA	CGC	CTA	CTC
AGG	AGT	ATG	ATT	CGG	CGT	CTG	CTT
GAA	GAC	GCA	GCC	TAA	TAC	TCA	TCC
GAG	GAT	GCG	GCT	TAG	TAT	TCG	TCT
GGA	GGC	GTA	GTC	TGA	TGC	TTA	TTC
GGG	GGT	GTC	GTT	TGG	TGT	TTG	TTT

70

**FIG. 7**

**FIG. 8****FIG. 9****FIG. 10**

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/14372

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 204/153.1, 400, 403; 435/6, 283.1, 287.1, 287.2; 436/518, 524, 807, 809

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAPLUS, MEDLINE, WPIDS, SCISEARCH, MEDLINE, EMBASE, BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,532,128 A (EGGERS et al) 02 July 1996. See entire document.	1-10
X	PEASE AC. Light-generated Oligonucleotide Arrays for Rapid DNA Sequence Analysis. Proc. Natl. Acad. Sci. May 1994. Vol 91. pages 5022-5026, especially Abstract.	1-10
X	STIMPSON DI. Real-time Detection of DNA Hybridization and Melting On Oligonucleotide Arrays by using Optical Wave Guides. Proc. Natl. Acad. Sci. July 1995. Vol 92. pages 6379-6383, especially Abstract	1-10
A,P	US 5,605,662 A (HELLER et al) 25 February 1997.	1-10

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

24 NOVEMBER 1997

Date of mailing of the international search report

23 DEC 1997

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

Telephone No. (703) 308-0196

**INTERNATIONAL SEARCH REPORT**International application No.  
PCT/US97/14372**C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	US 5,571,639 A (HUBBELL et al) 05 November 1996.	1-10

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/14372

A. CLASSIFICATION OF SUBJECT MATTER:  
IPC (6):

G01N 27/00, 27/26, 33/543, 33/551; C12Q 1/68; A01N 1/02; C12M 1/00;

A. CLASSIFICATION OF SUBJECT MATTER:  
US CL :

204/153.1, 400, 403; 435/6, 283.1, 287.1, 287.2; 436/518, 524, 807, 809